DOCUMENT RESUME

ED 460 217 UD 034 669

AUTHOR	Borman, Geoffrey D.; Hewes, Gina M.
TITLE	The Long-Term Effects and Cost-Effectiveness of Success for All.
INSTITUTION	Center for Research on the Education of Students Placed At Risk, Baltimore, MD.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
REPORT NO	CRESPAR-R-53
PUB DATE	2001-10-00
NOTE	39p.
CONTRACT	R-117-D40005
AVAILABLE FROM	Center for Research on the Education of Students Placed At Risk, Johns Hopkins University, 3003 North Charles Street, Suite 200, Baltimore, MD 21218. Tel: 410-516-8808; Fax: 410-516-8890; Web site: http://www.csos.jhu.edu.
PUB TYPE	Reports - Research (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Academic Achievement; Class Size; *Cost Effectiveness; *Early Intervention; *Early Reading; Elementary Secondary Education; High Risk Students; Program Effectiveness; *Reading Skills; Special Education; Urban Schools
IDENTIFIERS	Baltimore City Public Schools MD; Carolina Abecedarian Project NC; Perry Preschool MI; *Reform Efforts; Risk Reduction; *Success for All Program; Tennessee

ABSTRACT

This report presents results from an analysis of the long-term benefits and costs of Success for All (SFA), a school reform program that focuses on promoting early reading success among educationally at-risk students. The program was piloted and implemented in Baltimore elementary schools. This analysis tracks the educational outcomes through eighth grade for the original SFA students and for a quasi-experimental, untreated control group of students from matched comparison schools. The study also examines differential costs associated with SFA and control students' schooling through eighth grade. Finally, it compares the relative cost-effectiveness of SFA and three other initiatives (Perry Preschool, the Abecedarian Project, and the Tennessee class-size experiment). Data on the 1986-87 through 1998-99 school years were collected from computerized files provided by the Baltimore City Public School System. Results indicated that relative to controls, SFA students completed eighth grade at a younger age, at the same educational expense, and with better achievement outcomes, fewer special education placements, and less frequent retentions. Comparison to the other interventions suggests that SFA provides the strongest educational benefits for the dollar, though no single program may be relied upon as the "great equalizer." (Contains 45 references.) (SM)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 460 217

CRESPAR

THE LONG-TERM EFFECTS AND COST-EFFECTIVENESS OF SUCCESS FOR ALL

Geoffrey D. Borman
Gina M. Hewes

Report No. 53 / October 2001

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

FUNDED BY
OFFICE OF EDUCATIONAL RESEARCH AND IMPROVEMENT

JOHNS HOPKINS UNIVERSITY & HOWARD UNIVERSITY

CENTER FOR RESEARCH ON THE EDUCATION OF STUDENTS PLACED AT RISK

UD034669

BEST COPY AVAILABLE

THE LONG-TERM EFFECTS AND COST-EFFECTIVENESS OF SUCCESS FOR ALL

Geoffrey D. Borman
University of Wisconsin-Madison

Gina M. Hewes
Johns Hopkins University

Report No. 53

October 2001

This report was published by the Center for Research on the Education of Students Placed At Risk (CRESPAR), a national research and development center supported by a grant (No. R-117-D40005) from the Office of Educational Research and Improvement (OERI), U.S. Department of Education. The content or opinions expressed herein do not necessarily reflect the views of the Department of Education or any other agency of the U.S. Government. An on-line version of this report is available at our web site: *www.csos.jhu.edu*.

CRESPAR TECHNICAL REPORTS

1996

1. *The Talent Development High School: Essential Components*—V. LaPoint, W. Jordan, J.M. McPartland, D.P. Towns
2. *The Talent Development High School: Early Evidence of Impact on School Climate, Attendance, and Student Promotion*—J.M. McPartland, N. Legters, W. Jordan, E.L. McDill
3. *The Talent Development Middle School: Essential Components*—S. Madhere, D.J. Mac Iver
4. *The Talent Development Middle School: Creating a Motivational Climate Conducive to Talent Development in Middle Schools: Implementation and Effects of Student Team Reading*—D.J. Mac Iver, S.B. Plank
5. *Patterns of Urban Student Mobility and Local School Reform: Technical Report*— D. Kerbow
6. *Scaling Up: Lessons Learned in the Dissemination of Success for All*—R.E. Slavin, N.A. Madden
7. *School-Family-Community Partnerships and the Academic Achievement of African American, Urban Adolescents*—M.G. Sanders
8. *Asian American Students At Risk: A Literature Review*—S.-F. Siu

1997

9. *Reducing Talent Loss: The Impact of Information, Guidance, and Actions on Postsecondary Enrollment*—S.B. Plank, W.J. Jordan
10. *Effects of Bilingual Cooperative Integrated Reading and Composition on Students Transitioning from Spanish to English Reading*—M. Calderón, R. Hertz- Lazarowitz, G. Ivory, R.E. Slavin
11. *Effective Programs for Latino Students in Elementary and Middle Schools*—O.S. Fashola, R.E. Slavin, M. Calderón, R. Duran
12. *Detracking in a Racially Mixed Urban High School*—R. Cooper
13. *Building Effective School-Family-Community Partnerships in a Large Urban School District*—M.G. Sanders
14. *Volunteer Tutoring Programs: A Review of Research on Achievement Outcomes*— B.A. Wasik
15. *Working Together to Become Proficient Readers: Early Impact of the Talent Development Middle School's Student Team Literature Program*—D.J. Mac Iver, S.B. Plank, R. Balfanz
16. *Success for All: Exploring the Technical, Normative, Political, and Socio-Cultural Dimensions of Scaling Up*—R. Cooper, R.E. Slavin, N.A. Madden
17. *MathWings: Early Indicators of Effectiveness* —N.A. Madden, R.E. Slavin, K. Simons
18. *Parental Involvement in Students' Education During Middle School and High School* —S. Catsambis, J.E. Garland

1998

19. *Success for All/Éxito Para Todos: Effects on the Reading Achievement of Students Acquiring English*—R.E. Slavin, N.A. Madden
20. *Implementing a Highly Specified Curricular, Instructional, and Organizational School Design in a High-Poverty, Urban Elementary School: Three Year Results*—B. McHugh, S. Stringfield
21. *The Talent Development Middle School: An Elective Replacement Approach to Providing Extra Help in Math—The CATAMA Program (Computer and Team-Assisted Mathematics Acceleration)*—D.J. Mac Iver, R. Balfanz, S.B. Plank
22. *School-Family-Community Partnerships in Middle And High Schools: From Theory to Practice*—M.G. Sanders, J.L. Epstein
23. *Sources of Talent Loss Among High-Achieving Poor Students*—W.J. Jordan, S.B. Plank
24. *Review of Extended-Day and After-School Programs and Their Effectiveness*—O.S. Fashola
25. *Teachers' Appraisals of Talent Development Middle School Training, Materials, and Student Progress*—E. Useem
26. *Exploring the Dynamics of Resilience in an Elementary School*—S.M. Nettles, F.P. Robinson
27. *Expanding Knowledge of Parental Involvement in Secondary Education: Effects on High School Academic Success*—S. Catsambis
28. *Socio-Cultural and Within-School Factors That Effect the Quality of Implementation of School-Wide Programs*—R. Cooper

1999

29. *How Students Invest Their Time Out of School: Effects on School Engagement, Perceptions of Life Chances, and Achievement*—W.J. Jordan, S.M. Nettles
30. *Disseminating Success for All: Lessons for Policy and Practice*—R.E. Slavin, N.A. Madden
31. *Small Learning Communities Meet School-to-Work: Whole-School Restructuring for Urban Comprehensive High Schools*—N.E. Legters
32. *Family Partnerships with High Schools: The Parents' Perspective*—M.G. Sanders, J.L. Epstein, L. Connors-Tadros
33. *Grade Retention: Prevalence, Timing, and Effects*—N. Karweit
34. *Preparing Educators for School-Family-Community Partnerships: Results of a National Survey of Colleges and Universities*—J.L. Epstein, M.G. Sanders, L.A. Clark
35. *How Schools Choose Externally Developed Reform Designs*—A. Datnow
36. *Roots & Wings: Effects of Whole-School Reform on Student Achievement*—R.E. Slavin, N.A. Madden
37. *Teacher Collaboration in a Restructuring Urban High School*—N.E. Legters
38. *The Child First Authority After-School Program: A Descriptive Evaluation*—O.S. Fashola
39. *MathWings: Effects on Student Mathematics Performance*—N.A. Madden, R.E. Slavin, K. Simons
40. *Core Knowledge Curriculum: Three-Year Analysis of Implementation and Effects in Five Schools*—B. McHugh, S. Stringfield
41. *Success for All/Roots & Wings: Summary of Research on Achievement Outcomes*—R.E. Slavin, N.A. Madden

2000

42. *The Role of Cultural Factors in School Relevant Cognitive Functioning: Synthesis of Findings on Cultural Contexts, Cultural Orientations, and Individual Differences*—A.W. Boykin, C.T. Bailey
43. *The Role of Cultural Factors in School Relevant Cognitive Functioning: Description of Home Environmental Factors, Cultural Orientations, and Learning Preferences* —A.W. Boykin, C.T. Bailey
44. *Classroom Cultural Ecology: The Dynamics of Classroom Life in Schools Serving Low-Income African American Children*—C.M. Ellison, A.W. Boykin, D.P. Towns, A. Stokes
45. *An "Inside" Look at Success for All: A Qualitative Study of Implementation and Teaching and Learning*—A. Datnow, M. Castellano
46. *Lessons for Scaling Up: Evaluations of the Talent Development Middle School's Student Team Literature Program*—S.B. Plank, E. Young
47. *A Two-Way Bilingual Program: Promise, Practice, and Precautions*—M. Calderón, Argelia Carreón
48. *Four Models of School Improvement: Successes and Challenges in Reforming Low-Performing, High-Poverty Title I Schools*—G.D. Borman, L. Rachuba, A. Datnow, M. Alberg, M. Mac Iver, S. Stringfield, S. Ross
49. *National Evaluation of Core Knowledge Sequence Implementation: Final Report*— S. Stringfield, A. Datnow, G. Borman, L. Rachuba
50. *Core Knowledge Curriculum: Five-Year Analysis of Implementation and Effects in Five Maryland Schools*—M.A. Mac Iver, S. Stringfield, B. McHugh

2001

51. *Effects of Success for All on TAAS Reading: A Texas Statewide Evaluation*— E.A. Hurley, A. Chamberlain, R.E. Slavin, N.A. Madden
52. *Academic Success Among Poor and Minority Students: An Analysis of Competing Models of School Effects*—G.D. Borman, L.T. Rachuba
53. *Thee Long-Term Effects and Cost-Effectiveness of Success for All*—G.D. Borman, G.M. Hewes
54. *Neighborhood and School Influences on the Family Life and Mathematics Performance of Eighth-Grade Students*—S. Catsambis, A.A. Beveridge
55. *The Public School Superintendency in the 21st Century: The Quest to Define Effective Leadership*—J.Y. Thomas

PLEASE NOTE

Reports are available from:

Publications Department
Attn: Diane Diggs
CRESPAR/Johns Hopkins University
3003 N. Charles Street, Suite 200
Baltimore MD 21218
410-516-8808 phone/410-516-8890 fax
www.csos.jhu.edu

BEST COPY AVAILABLE

THE CENTER

Every child has the capacity to succeed in school and in life. Yet far too many children, especially those from poor and minority families, are placed at risk by school practices that are based on a sorting paradigm in which some students receive high-expectations instruction while the rest are relegated to lower quality education and lower quality futures. The sorting perspective must be replaced by a "talent development" model that asserts that all children are capable of succeeding in a rich and demanding curriculum with appropriate assistance and support.

The mission of the Center for Research on the Education of Students Placed At Risk (CRESPAR) is to conduct the research, development, evaluation, and dissemination needed to transform schooling for students placed at risk. The work of the Center is guided by three central themes—ensuring the success of all students at key development points, building on students' personal and cultural assets, and scaling up effective programs—and conducted through research and development programs in the areas of early and elementary studies; middle and high school studies; school, family, and community partnerships; and systemic supports for school reform, as well as a program of institutional activities.

CRESPAR is organized as a partnership of Johns Hopkins University and Howard University, and supported by the National Institute on the Education of At-Risk Students (At-Risk Institute), one of five institutes created by the Educational Research, Development, Dissemination and Improvement Act of 1994 and located within the Office of Educational Research and Improvement (OERI) at the U.S. Department of Education. The At-Risk Institute supports a range of research and development activities designed to improve the education of students at risk of educational failure because of limited English proficiency, poverty, race, geographic location, or economic disadvantage.

ABSTRACT

A few renowned early interventions have compelling evidence of enduring achievement effects for at-risk children—Perry Preschool, the Abecedarian Project, and the Tennessee class-size experiment. The costs and potential for national dissemination of such model programs, though, represent key practical concerns. This study examines the long-term outcomes and costs of another popular early intervention—Success for All. Relative to controls, Success for All students complete eighth grade at a younger age, with better achievement outcomes, fewer special education placements, less frequent retentions, and at the same educational expense. Further comparisons to the three prominent interventions suggest that Success for All provides the strongest educational benefits for the dollar. However, no single program may be relied on as the "great equalizer."

ACKNOWLEDGMENTS

T!.e authors thank Spyros Konstantopoulos, Hank Levin, Nancy Madden, Stephen Meyer, Allan Odden, and Bob Slavin for their helpful feedback on previous drafts of this report.

INTRODUCTION

Few American policies for improving educational equality and productivity have convincing evidence of long-lasting benefits to students and society. Although a number of early interventions show moderate to strong immediate impacts, these effects generally fade away after several years, thus leaving program participants with academic outcomes no different than those of nonparticipants. This has been a common criticism of compensatory education initiatives, such as Head Start (Lee, Brooks-Gunn, & Schnur, 1988; McKey, Condelli, Ganson, Barret, McConkey, & Plantz, 1985) and Title I (Borman & D'Agostino, 2001; Carter, 1984), and many other early childhood interventions (White, 1985).

Along with the programs themselves, the supporting research literature also has been criticized. Barnett (1995) concluded that studies of only two such programs provided compelling data showing long-term positive effects on outcomes, such as achievement, grade retentions, special education placements, high school graduation, and socialization: the Abecedarian Project and Perry Preschool studies. More recently, a longitudinal followup to another well-designed study, the Tennessee Student/Teacher Achievement Ratio (STAR) experiment, provided strong support for the long-term efficacy of class-size reductions in the early elementary school years (Finn, Gerber, Achilles, & Boyd-Zaharias, 2001; Nye, Hedges, & Konstantopolous, 2000). These three studies provide important evidence of the sustained effects of early educational policies and programs on students' academic outcomes through middle school and, in some cases, into adulthood. Moreover, they inspire strong hope for the American ideal for education, as expressed by Horace Mann, to be the "great equalizer," or "balance wheel of the social machinery."

From a more mundane perspective, policymakers and practitioners also are interested in the feasibility and costs associated with replication of these model programs and their outcomes. A frequently voiced concern about Head Start and other state-funded programs that have been inspired by the Perry Preschool and Abecedarian efforts is whether they have captured the seemingly magical essence of the earlier demonstration projects (Gomby, Larner, Stevenson, Lewit, & Behrman, 1995). Similarly, the implementations and outcomes of class-size reduction initiatives that have emerged in response to the STAR findings, including the prominent $1.5 billion per year California effort, have not necessarily matched the results of the small and highly controlled Tennessee experiment (Bohrnstedt & Stecher, 1999). Despite these apparently sobering outcomes regarding replication, there is evidence suggesting that the long-term benefits to society of the well-implemented pilot models, most notably the Perry Preschool program, considerably outweigh their substantial costs (Barnett, 1985; Schweinhart & Weikart, 1986).

In this report, we present the results from the first analysis of the long-term benefits and costs of another important and widely replicated early intervention—Success for All

(SFA). Currently implemented in approximately 2,000 schools serving more than million children throughout the United States, Success for All is a school reform program that focuses on promoting early reading success among educationally at-risk students. The program was developed by Robert Slavin, Nancy Madden, and colleagues at the request of the Baltimore City Public School System, and was piloted in one Baltimore elementary school during the 1987-88 school year. Four additional Baltimore schools implemented Success for All during the subsequent year, 1988-89. Our analyses track the educational outcomes through the eighth grade for the original Success for All students and for a quasi-experimental untreated control group composed of students from matched comparison schools. In addition, we examine the differential costs associated with Success for All and control students' schooling through eighth grade. We conclude by comparing the relative cost-effectiveness of Success for All, Perry Preschool, the Abecedarian Project, and the Tennessee class-size reduction initiative.

A Conceptual Framework for Understanding the Enduring Effects of Early Interventions

What features of early interventions are related to long-lasting cognitive benefits and why would we expect to find enduring effects from them? A review by Ramey and Ramey (1998) of the major findings of rigorous studies provides consistent answers which help explain the enduring effects of the Perry Preschool program, Abecedarian Project, and Tennessee class-size reduction effort. The conceptual framework, *biosocial developmental contexualism*, derived from this review predicts that fragmented, weak efforts in early intervention are not likely to succeed, whereas intensive, high-quality ecologically pervasive interventions can and do. This framework highlights six principles: developmental timing; program intensity; direct provision of learning experiences; program breadth and flexibility; individual differences in program benefits; and environmental maintenance of development. These six principles, along with select references to how the three model early interventions serve as exemplars of these principles, are discussed below.

The work of Bloom (1964), among others, stresses the great malleability of human development in the early years of life. The principle of developmental timing supports the notion that interventions which begin earlier are especially advantageous in that they may capitalize on this malleability and help alter at-risk children's long-range developmental trajectories before they have had the chance to diverge substantially from the trajectories of more advantaged children. The Abecedarian Project and other early education programs showing some of the largest effects enrolled children during infancy and continued for several years. The Perry Preschool and Tennessee STAR interventions also started early, serving children at the ages of 3 to 4 and 5, respectively. Although the optimal timing for early intervention is open to some debate, the sustained effects on reading achievement

through mid-adolescence for the earlier interventions, Abecedarian, g=0.53 (Ramey, Campbell, Burchinal, Skinner, Gardner, & Ramey, 2000) and Perry Preschool, g=0.51 (Scheinhart, Barnes, & Weikart, 1993) were larger than the two effect estimates for STAR's reductions in kindergarten through third grade class sizes, g=0.22 (Finn et al., 2001), and g=0.32 (Nye et al., 2000).[1]

Ramey and Ramey (1998) also point out that early interventions that provide highly intensive services (indexed by variables including the number of hours per day, days per week, and weeks per year that the intervention is offered) produce greater sustained effects than do less intensive programs. Perry Preschool and Abecedarian provided a range of intensive services for both children and parents. Similarly, those children from the Tennessee STAR study with more years of exposure to reduced class sizes showed stronger sustained effects than students with fewer years in a small class (Finn et al., 2001; Nye et al., 2000). All three interventions also serve as prime examples of Ramey and Ramey's (1998) third principle, in that they provided direct educational experiences to participating students rather than relying exclusively on intermediary routes (e.g., parent training or home visits) to change children's competencies.

The fourth principle suggests that those interventions which offer a comprehensive approach, including a strong educational program, social services, family support, and individualized assistance, tend to produce larger effects than do interventions with a narrower focus. The multipronged Perry Preschool and Abecedarian programs, which offered arrays of services to both children and parents, provide strong prototypes of this principle. Although no data was collected in Tennessee to document how reduced class sizes affected teachers' interactions with children, there is other evidence suggesting that smaller classes encourage more individualized attention to students' academic and personal problems (Stasz & Stecher, 2000). Furthermore, findings from STAR revealed that children who attended small classes achieved improved social and behavioral outcomes in comparison to their counterparts in regular-sized classes (Finn, Fulton, Zaharias, & Nye, 1989).

The benefits accruing to students from these interventions tend to be driven by the principle of individual differences in program benefits. Some children show greater benefits from participation than others and these differences appear to be related to aspects of children's initial risk conditions. An important example of this principle comes from the Tennessee STAR study, which documented that minority students typically benefitted more from reductions in class size than non-minority children (Finn & Achilles, 1999).

A final caveat affecting long-term outcomes is suggested by Ramey and Ramey's (1998) sixth principle, which suggests that the initial positive effects of early intervention will diminish over time to the extent that there are not adequate follow-ups to maintain children's social, behavioral, and academic gains. The only intervention that assigned some students to receive follow-up services, the Abecedarian Project, showed that children

receiving a longer duration of services, during both preschool and elementary school, achieved better long-term reading outcomes than students with shorter durations of participation (Campbell & Ramey, 1995). Over time, though, most analyses of these three programs showed slighter effects after the interventions had ceased.

Success for All

Success for All is, arguably, the most widely implemented, widely researched, and widely critiqued educational program in the United States. Today, schools may purchase the program from the not-for-profit Success for All Foundation as a comprehensive package, which includes materials, training, ongoing professional development, and a highly specified "blueprint" for implementing and sustaining the model. Schools that elect to adopt Success for All implement a schoolwide program that organizes resources to attempt to ensure that every child will be successful in reading from the beginning of their time in school. Rather than special classes, retentions in grade, and other forms of remediation, the program emphasizes prevention and early, intensive intervention designed to detect and resolve reading problems as early as possible, before they become serious.

Schools that adopt Success for All implement a series of reading programs, beginning as early as prekindergarten and extending through the later elementary grades. Students spend most of their day in traditional, age-grouped classes, but are regrouped across grades for reading lessons targeted to specific reading levels. Teachers assess each student's reading performance at eight-week intervals and make regrouping changes based on the results. Rather than being placed in special classes or retained in grade, students who need additional help receive one-on-one tutoring to get them back on track. A Success for All school also establishes a Family Support Team, serving to increase parents' participation in school generally and to identify and address particular problems such as irregular attendance, vision correction, or problems at home. Finally, each Success for All school designates a full-time Program Facilitator who oversees the daily operation of SFA, provides assistance where needed, and coordinates the various components of the program. These are the main features of Success for All, both as originally conceived and as currently disseminated.[2]

Similar to the Abecedarian Project, Perry Preschool, and Tennessee class-size reduction initiatives, Success for All is relatively costly. A recent review of 26 school reform models identified Success for All as one of the most expensive reforms, with estimates of first-year personnel, materials, and training costs of between $70,000 and $270,000 for a typical school (Herman, Aladjem, McMahon, Masem, Mulligan, O'Malley, Quinones, Reeve, & Woodruff, 1999). In estimating the costs of implementing three reforms, Success for All, the Accelerated Schools model, and the School Development Program, King (1994) also concluded that the per-school costs for Success of All, which ranged from $261,060 to $646,500 per year, were the highest. However, the Success for All developers argue that

schools with high enrollments of poor children generally have sufficient supplemental federal and state Title I dollars to implement a credible form of the model, and these resources often are further augmented by reallocated funds and personnel from special education, desegregation settlements, and other sources (Slavin, Madden, Dolan, Wasik, Ross, & Smith, 1994).

Previous studies of the program's outcomes have focused primarily on its immediate achievement effects. The prototypical Success for All evaluation has used a quasi-experimental, untreated control group design employing matched student samples from two similar schools. Although there are exceptions (cf Jones, Gottfredson, & Gottfredson, 1997), the clear majority of these studies has documented positive achievement outcomes and reduced retention and special education placement rates for Success for All students (Slavin & Madden, 2001). Previous analyses of the short-term effectiveness of the program in the five Baltimore elementary schools revealed effect sizes above 0.50 on individually administered reading tests in grades 1 through 3 (Madden, Slavin, Karweit, Dolan, & Wasik, 1993). The most educationally at-risk students, who were identified as the lowest-scoring 25% on the pretest, showed the strongest effects from the intervention, with effect sizes at or close to 1.00 (Madden et al., 1993). These outcomes from the pilot Success for All schools, and their successful replication in numerous other schools, are impressive. Nevertheless, the long-term outcomes of the intervention remain untested and the overall cost-effectiveness of the program is open to debate.

Hypotheses of the Current Study

We hypothesized that, similar to the studies of the Perry Preschool program, the Abecedarian Project, and the Tennessee class-size reduction effort, we would find sustained effects on important academic outcomes for students who had participated in the Success for All program. Further, although Success for All is a somewhat costly intervention, we expected that the model's focus on prevention would not be substantially more expensive than the traditional emphasis on remediation, primarily in the form of retention and special education, to which control students would, most likely, be exposed.

Why would we expect Success for All to promote long-term positive effects for students? The robust short-term effects for the program present a compelling empirical case to support this expectation, but the application of Ramey and Ramey's (1998) six-principle framework provides an added theoretical basis. First, Success for All stresses early intervention and prevention, with services beginning in prekindergarten or kindergarten. Second, the program stresses intensity, from its daily 90-minute reading periods to its multi-year approach. Third, learning experiences are delivered to students directly, efficiently, and effectively in regrouped classrooms and in one-on-one tutoring sessions. Fourth, program breadth and flexibility are exemplified by the range of services offered to parents and to

students by the Family Support Team, classroom teachers, and tutors. With respect to the fifth principle, individual differences in program benefits, Success for All appears to have the most pronounced impacts on the students with the greatest academic needs. The current study helps us examine Ramey and Ramey's (1998) sixth principle: environmental maintenance of development. That is, after discontinuation of the Success for All intervention during the later elementary grades, do children maintain their academic gains through middle school into mid-adolescence and, if so, at what cost?

METHOD

Data

All data, from the 1986-87 through 1998-99 school years, were abstracted from computerized files provided by the Baltimore City Public School System (BCPSS). All student background, transcript, and achievement data were collected by the BCPSS as part of its annual district-wide testing and data collection programs. We utilized two yearly files, the Pupil Information File (PIF) which contains basic data such as race/ethnicity, gender, grade level, school, and special education participation information, and a data file containing student-level standardized test results. We used data from the district-administered California Achievement Test (CAT) and Comprehensive Test of Basic Skills, Fourth Edition (CTBS/4). The CAT provided the pretest reading data used in our analyses and the CTBS/4 provided our eighth-grade reading and math achievement outcomes.

Sample

The sample included students from the original Success for All elementary schools from Baltimore, Maryland—Abbottston, City Springs, Dr. Bernard Harris, Harriet Tubman, and Dallas F. Nicholas, Sr.—and their five matched control schools, respectively Tench Tilghman, Collington Square, Patapsco, Harlem Park, and Charles Carroll of Carrollton. These schools were matched on the demographic characteristics of their students and received similar base levels of funding. Therefore, the only clear difference between the Success for All and matched control schools was the added marginal resources associated with the Success for All implementation.[3] Beginning with the baseline year of Success for All implementation, which was 1987-88 for Abbottston and 1988-89 for the other four Success for All schools, we identified four independent cohorts of first-grade students from the 1987-88, 1988-89, 1989-90, and 1990-91 academic years. Combining these four cohorts of students yielded a total sample of 1,388 Success for All and 1,848 control students.

These samples were reduced in our analyses due to listwise deletion of cases with missing data on the background variables and outcomes. For analysis of eighth-grade test score outcomes, the final analytical sample sizes were 581 and 729 for Success for All and control students, respectively. All other analyses based on transcript data from the PIF included 735 Success for All and 995 control students. Listwise deletion of cases with missing values did not cause differential attrition rates by program condition, leaving 42% of the baseline sample of 1,388 Success for All students and 39% of the 1,848 baseline controls for our analyses of eighth-grade test score outcomes $\chi^2(1, N=1310)=1.91$ and 53% and 54% of the respective baseline Success for All and control samples for our analyses of transcript outcomes $\chi^2(1, N=1730)=0.14$.[4]

We also identified subsamples of the lowest achieving students for analysis. We defined the low-achieving subsamples as those students whose kindergarten pretest scores fell in the bottom 25% of their school for that year. From the sample of students with valid background data and eighth-grade test score results, we identified 148 Success for All and 184 control low achievers. The final low-achieving subsamples of students with valid background and eighth-grade test data were 188 and 250 for Success for All and control students, respectively.

Baseline data for the analytical samples are displayed in Table 1. Sample sizes, means, and standard deviations are provided for the full sample and low-achieving sample and for the samples used in the achievement and transcript analyses. The samples were comprised of students who were overwhelmingly poor, as indicated by free or reduced-price lunch eligibility, and African American. Students' baseline ages, recorded during September of first grade, were relatively typical, as were the gender splits.

We found consistent differences between the Success for All and control group students on only one background characteristic: reading pretest. In all analytical samples, the differences favored control students by more than one quarter of one standard deviation. Attempts to match Success for All and control students one-on-one failed, as statistically significant pretest differences remained in the optimally matched samples.[5] Therefore, attempts to analyze the one-on-one matched samples were not pursued and, instead, all analyses included the full control and Success for All samples while statistically controlling for pretest differences. In addition, in the case of one analytical sample, the low-achieving sample used for the PIF transcript analyses, we found one relatively small but statistically significant difference between Success for All and control group students on baseline age. Analyses of the transcript outcomes for low achievers, thus, also included baseline age as a covariate.

Table 1
Baseline Data for Analytical Samples

	Success for All			Control		
	N	*M*	*SD*	*N*	*M*	*SD*
Full Sample						
Achievement Analysis	581			729		
CAT Reading Pretest		-0.04***	0.99		0.29	0.84
Female		0.56	0.50		0.56	0.50
African American		0.99	0.07		0.99	0.10
Free/Reduced Lunch		0.91	0.29		0.90	0.29
Age		6.27	0.39		6.28	0.36
Transcript Analysis	738			995		
CAT Reading Pretest		-0.17***	1.04		0.17	0.91
Female		0.50	0.50		0.53	0.50
African American		0.99	0.12		0.99	0.08
Free/Reduced Lunch		0.92	0.27		0.91	0.29
Age		6.30	0.42		6.32	0.40
Low-Achieving Sample						
Achievement Analysis	148			184		
CAT Reading Pretest		-1.28***	0.60		-0.81	0.55
Female		0.55	0.50		0.49	0.50
African American		0.98	0.14		0.99	0.07
Free/Reduced Lunch		0.92	0.27		0.90	0.30
Age		6.25	0.40		6.29	0.44
Transcript Analysis	188			250		
CAT Reading Pretest		-1.46***	0.60		-1.04	0.57
Female		0.45	0.50		0.48	0.50
African American		0.98	0.14		0.99	0.09
Free/Reduced Lunch		0.94	0.25		0.90	0.30
Age		6.26*	0.41		6.35	0.46

Measures

Student Outcomes. Success for All is designed specifically to affect three outcomes: achievement, grade-level progression (or retention), and special education placements. Our analyses included measures of all three outcomes through mid-adolescence. We used Total Reading and Total Math scale scores from the district-wide administration of the CTBS/4 during the fall of eighth grade. Because this is the only time at which all Baltimore City students are tested after elementary school, we took a student's CTBS/4 score from when he or she first started eighth grade, regardless of when or at what age a student reached eighth grade.

Our analyses of special education placements and retentions in grade included four variables: the number of years a student was placed in special education during elementary school; the number of years a student was placed in special education during middle school; the number of times a student was retained during elementary school; and the number of times a student was retained during middle school. Retentions and special education placements during elementary school were assumed to be affected directly by Success for All policies, which explicitly aim to limit these outcomes, and indirectly through students' improved academic outcomes and decreased needs for special services. In contrast, middle school students' retentions and special education placements occur after leaving the Success for All elementary school program and, thus, are not directly affected by explicit Success for All policies.

Pretest Covariate. The CAT pretest was administered during the spring prior to each student's cohort year. For the 1988, 1989, and 1990 cohorts, pretest was defined as the student's CAT Total Reading score. Because a Total Reading score was not available for most students from the 1991 cohort, we derived a similar score by taking the mean of three CAT reading subscales, Phonics Analysis, Structural Analysis, and Vocabulary (α=.88). However, this resulted in a mean pretest scale score for the 1991 cohort that was considerably higher than the mean pretest scores for the other three cohorts. Therefore, we standardized the pretest scores within each cohort by converting them to z-scores. It is the z-score of the CAT pretest that is used in all analyses.

Cost Estimates. Our second research question concerned the relative costs of schooling for Success for All and control students. Several cost estimates were required for this comparison. First, we established the costs of Success for All using the ingredients method (Levin & McEwan, 2001). Summaries of the total and marginal costs derived from this analysis are summarized by school and by year in Table 2. Detailed descriptions of the program ingredients were provided by Slavin, Madden, Karweit, Dolan, and Wasik (1992). The authors reported the numbers of tutors, whether Program Facilitators were full-time or half-time employees, and additional support staff employed at each school. We used these

data to determine the number of personnel required for the Success for All program at each school.

We derived information regarding the current market values for personnel salaries from the U.S. Department of Labor, Bureau of Labor Statistics' (1999) 1998 Occupational Employment Statistics (OES) national survey. Costs for Success for All tutors were estimated using the OES reported median annual salary for teachers ($36,110), the costs for the lead Family Support Team member were based on the OES median salary for counselors ($38,650), and costs for additional Family Support Staff were based on the OES median salary for teacher aides ($16,280). To express all salaries in current dollar amounts, all of these 1998 salaries were then adjusted to constant 2000 dollars using gross domestic product implicit price deflators. Finally, we calculated the additional costs of fringe benefits as 26.35% of instructional staff salary and 27.99% of support staff salary for each of these two categories of personnel (U.S. Bureau of the Census, 1997).

In addition to personnel costs, implementation of the Success for All program at the Baltimore schools involved expenses for training, materials, and ongoing professional development. Current market value estimates of the costs of these goods and services were derived from pricing information available on the Success for All internet site, www.successforall.net: $70,000 to $85,000 for the first year of implementation; $26,000 to $30,000 for the second year; and $23,000 to $25,000 for the third and later years. The midpoint of each of these ranges was used as the estimate for the Baltimore schools.

The Success for All programs operating at the five Baltimore schools focused on grades 1 through 5. Consequently, the school-by-school and year-by-year marginal costs listed in Table 2 were divided by each school's yearly student enrollment in grades 1 through 5 to obtain the yearly per-pupil costs of Success for All used in our analysis. We assigned these school-specific and year-specific per-pupil costs to all students enrolled in grades 1 through 5 in a Success for All school. These marginal cost estimates served as the core of the cost analyses of Success for All and also provide valuable information regarding the current expenses associated with replicating the Success for All interventions as originally configured in the five Baltimore schools.[6]

Two additional estimates were needed to calculate the cost of each student's education. First, as a measure of the overall market value of each child's basic education program, we used the annual average per-pupil expenditure for the United States, which was $5,330 for the 1998-99 school year (Tom Parrish, personal communication, February 15, 2001), or $5,423 in constant 2000 dollars. Second, as an estimate of the overall market value of special education services, we utilized the annual average additional per-pupil expense of special education, which was $6,404 in 1998-99 (Tom Parrish, personal communication, February 15, 2001), or $6,515 after adjusting to 2000 dollars.

Based on yearly data from each grade level, 1 through 8, we estimated the costs associated with each year of a student's educational program as a combination of three possible costs: (a) the current base per-pupil expense of $5,423; (b) the additional per-pupil cost of special education, $6,515 and; (c) the school-specific and year-specific per-pupil cost of Success for All. The sum of each student's yearly cost of schooling through the end of middle school, that is, through the successful completion of eighth grade, was the final figure that we analyzed as an outcome variable. These analyses were restricted to those students for whom we were able to establish yearly information regarding special education assignments, school assignments, and grade-level progression from grades 1 through 8.

Analytical Approach

The primary inferential method that we used was standard analysis of covariance (ANCOVA), controlling for pretest score. In one case, for our analysis of low-achieving students' transcript outcomes, we also controlled for a statistically significant Success for All-control difference for baseline age. Each adjusted mean difference between Success for All and control students that we obtained from the ANCOVAs was divided, or standardized, by the pooled posttest standard deviation for the outcome. The resulting standardized differences, or effect sizes, provide summaries of the magnitude of each effect and are interpretable as the number of standard deviation units separating Success for All students from control students on the covariate-adjusted outcomes.

RESULTS

Table 3 displays the results for the analyses of achievement outcomes and transcript outcomes for both the total sample and the low-achieving sample. All statistically significant Success for All-control differences revealed by the ANCOVAs are indicated in the far-right column displaying the pretest-adjusted effect size values. After controlling for the kindergarten pretest differences, Success for All students had higher eighth grade CTBS/4 reading and math scale scores, than did control students, Fs (1, 1307)=29.84 and 4.40, $ps<$ .001 and .05, respectively. Expressing the pretest-adjusted eighth grade test scores in grade equivalents, Success for All students held a six-month advantage over control students in reading, 5.7 versus 5.1, and a three-month advantage in math, 6.2 versus 5.9. The magnitude of the Success for All reading effect size exceeded one quarter of one standard deviation. These results suggest that the achievement effects of the program, especially in its target subject area of reading, attained both statistical and practical significance.

Table 2
Program Ingredients and Total and Marginal Costs of Success for All by School and by Year

	Abbottston	City Springs	Dallas Nicholas	Dr. Bernard Harris	Harriet Tubman
Certified Reading Tutors FTEs	6	9	2	3	3
Salary per Certified Reading Tutor	$37,291.00	$37,291.00	$37,291.00	$37,291.00	$37,291.00
Benefits per Certified Reading Tutor	9,826.18	9,826.18	9,826.18	9,826.18	9,826.18
Tutors Sub-Total:	$282,703.08	$424,054.62	$94,234.36	$168,816.36	$168,816.36
Family Support Staff FTEs	2	2.5	0	0	0
Salary per Social Worker	$30,940.00	$30,940.00			
Benefits per Social Worker	8,660.11	8,660.11			
Salary per Family Support Staff	16,812.00	16,812.00			
Benefits per Family Support Staff	4,705.68	4,705.68			
Family Support Staff Sub-Total:	$61,117.79	$71,876.63			
Program Facilitator FTEs	1	1	0.5	0.5	0.5
Salary per Facilitator	$37,291.00	$37,291.00	$18,645.50	$18,645.50	$18,645.50
Benefits per Facilitator	9,826.18	9,826.18	4,913.09	4,913.09	4,913.09
Facilitator Sub-Total:	$47,117.18	$47,117.18	$23,558.59	$23,558.59	$23,558.59
Training, Implementation, Materials, Year 1	$77,500.00	$77,500.00	$77,500.00	$77,500.00	$77,500.00
Training, Implementation, Materials, Year 2	$28,000.00	$28,000.00	$28,000.00	$28,000.00	$28,000.00
Training, Implementation, Materials, Year 3	$24,000.00	$24,000.00	$24,000.00	$24,000.00	$24,000.00
Total Cost, Year 1:	$468,438.05	$620,548.42	$195,292.95	$269,874.95	$269,874.95
Marginal Cost, Year 1:	$468,438.05	$620,548.42	$124,617.18	$124,617.18	$124,617.18
Total Cost, Year 2:	$418,938.05	$571,048.42	$145,792.95	$220,374.95	$220,374.95
Marginal Cost, Year 2:	$418,938.05	$571,048.42	$75,117.18	$75,117.18	$75,117.18
Total Cost, Years 3+:	$414,938.05	$567,048.42	$141,792.95	$216,374.95	$216,374.95
Marginal Cost, Years 3+:	$414,938.05	$567,048.42	$71,117.18	$71,117.18	$71,117.18

Table 3
Achievement and Transcript Outcomes through Eighth Grade

	Success for All M^a	Control M^a	Pooled *SD*	*g*
Full Sample				
Achievement Outcomes				
CTBS/4 Total Reading scale score	716.97	703.71	46.42	0.29***
CTBS/4 Total Math scale score	723.29	718.08	47.13	0.11*
Transcript Outcomes				
Retained in elementary school	0.09	0.25	0.41	-0.39***
Retained in middle school	0.13	0.15	0.40	-0.04
Special education in elementary school	0.55	0.82	1.48	-0.18***
Special education in middle school	0.49	0.70	1.16	-0.18***
Age	14.20	14.33	0.58	-0.22***
Educational expenditures	53506.36	54766.47	16786.30	-0.08
Low-Achieving Sample				
Achievement Outcomes				
CTBS/4 Total Reading scale score	696.06	679.87	44.98	0.36**
CTBS/4 Total Math scale score	703.20	694.69	44.77	0.19
Transcript Outcomes				
Retained in elementary school	0.18	0.49	0.53	-0.59***
Retained in middle school	0.15	0.21	0.45	-0.15
Special education in elementary school	1.24	1.74	1.95	-0.26*
Special education in middle school	1.01	1.39	1.42	-0.27**
Age	14.29	14.52	0.66	-0.35***
Educational expenditures	61663.09	66246.71	21121.65	-0.22*

Note: * $p < .05$; ** $p < .01$; *** $p < .001$.

[a]Success for All and control mean columns display covariate-adjusted means.

The long-term special education outcomes in Table 3 also show statistically significant differences between Success for All students and control students, with Success for All students spending fewer years than control students enrolled in special education during the elementary school grades and middle school grades, *F*s (1, 1730)=15.43 and 16.59, respectively, *p*s<.001. After controlling for the pretest, Success for All students, on average, spent about half of 1 academic year (0.55) during elementary school in special education, compared to more than three quarters of a year (0.82) for control students. This result is not surprising in that an explicit policy of Success for All is to refrain from special education referrals, except under unusual circumstances. Therefore, the special education differences between Success for All and control students in elementary school may be interpreted as a result of this policy, the improved academic performance of Success for All students, or, most likely, some combination of both.

Special education placements after the elementary grades, in middle school, are not open to this combination of interpretations. In fact, if it were the case that the Success for All policy of reducing special education referrals were ill-advised, we might expect former Success for All students to have higher rates of special education placement after they left the program, on the theory that some of the students that were not referred to special education under Success for All actually were in need of referral. On the contrary, we see that Success for All students continued to be less likely than control students to participate in special education in the middle school grades. After controlling for the pretest, the average Success for All student spent approximately half of 1 school year (0.49) during middle school in special education versus more than two thirds of a school year (0.70) for the typical control student.

The findings for retentions in elementary and middle school are consistent with the special education findings, and bear the same importance for judging the effectiveness of Success for All. During the elementary grades, the typical Success for All student was retained close to 0 times (0.09), compared to control students' 0.25 retentions, $F(1, 1730) = 67.50$, $p<.001$. In middle school, though, Success for All and control students were retained at a statistically equivalent frequency, 0.13 and 0.15 times, respectively, $F(1, 1730)=0.52$. Like the outcomes for special education, Success for All students' lower number of retentions during the elementary school years most likely reflects a combination of both overt Success for All policy and improved student outcomes. Regarding the middle school outcomes, again, we might expect to see a higher frequency of retentions for Success for All students than for control students if the policy simply served to promote students who should have been retained. Instead, the results in Table 3 show that there is no difference in retention frequency for Success for All and control students during the middle school grades. Associated with the less frequent retentions of Success for All students, we also found that they successfully completed eighth grade at a younger age than control students, $F(1, 1730) = 20.94$, $p<.001$. The difference between the pretest-adjusted means indicated that Success

for All students, on average, finished eighth grade 1.5 months ahead of their control group peers.

Low Achievers

We repeated the above analyses on the subsample of initially low-achieving students scoring in the bottom 25% of their respective schools on the pretest. The results are shown at the bottom of Table 3. Consistent with previous analyses of the short-term outcomes for the Baltimore Success for All students (Madden et al., 1993; Slavin, Madden, Karweit, Livermon, & Dolan, 1990), and consistent with Ramey and Ramey's (1998) principle of individual differences in program benefits, the effect size magnitudes for all outcomes for the low-achieving sample exceeded those found for the full sample. Despite the smaller sample sizes and corresponding loss of statistical power, six of the eight outcomes achieved conventional levels of statistical significance.

For CTBS/4 reading, initially low-achieving Success for All students had a higher adjusted mean scale score than did initially low-achieving control students, $F(1, 329)=9.35$, $p<.01$. Expressing the covariate-adjusted outcomes in the grade equivalent metric, Success for All students were 7 months ahead of controls, with group means of 4.8 and 4.1, respectively. Although the Success for All students' eighth-grade math scores were nearly one fifth of one standard deviation higher than control students' outcomes, the difference did not attain statistical significance, $F(1, 329)=2.59$.

After adjusting for the pretest and baseline age for our analyses of the low-achievers' transcript outcomes, we found that the Success for All students spent fewer elementary school years than control students in special education and were retained less often, Fs (1, 435)=7.43 and 31.31, ps<.01 and .001, respectively. The special education and retention outcomes during the middle school years also favored Success for All students. The difference for special education placements was statistically significant, $F(1, 435)=8.62$, $p<.01$, but the difference for the retention outcome was not, $F(1, 435)=1.91$. The fewer overall retentions among Success for All students helped contribute to a statistically significant age advantage at the completion of eighth grade, $F(1, 435)=19.56$, $p<.001$. That is, after adjusting for the pretest, Success for All students completed the eighth grade 2.8 months earlier than did control students. These outcomes have all of the same implications discussed previously for the full sample.

The Cost of SFA

Our final transcript-based outcome is the estimated per-pupil costs of schooling through the end of middle school, or through the successful completion of eighth grade. The adjusted

average costs from the ANCOVA for both the total sample and low-achieving sample are displayed in Table 3.[7] These estimates include base per-pupil expenditures, special education costs, and school- and year-specific Success for All costs for each year of the students' schooling. After adjusting for the pretest, the costs associated with Success for All students' schooling through eighth grade were statistically equivalent to the costs of control students' schooling, F (1, 1730)=2.76. The mean unadjusted cost of Success for All students' schooling was $54,893.59 ($SD$=16346.70) and the mean unadjusted cost for control students was $53,737.55 ($SD$=17096.58).

The costs of schooling for low-achieving students were considerably higher, with an unadjusted mean of $63,580.42 ($SD$=20886.57) for Success for All students and an unadjusted mean of $64,804.87 ($SD$=21323.36) for control students. Controlling for the pretest and baseline age for our analyses of the low-achievers, the adjusted cost for Success for All students was lower than the cost for control students, F (1, 435)=4.69, p<.05. The magnitude of the difference, which was equivalent to an effect size of -0.22, suggests that the preventative approach of Success for All was considerably less costly than the remedial approach of more frequent special education placements and retentions, which was characteristic of the low-achieving control group.

DISCUSSION

The findings from this study have important implications for educational policy, theory, and practice. From a policy perspective, our results indicate that a nationally disseminated elementary school program may deliver enduring educational benefits to the students it serves at no additional cost. Specifically, Success for All students complete eighth grade at a younger age, with better achievement outcomes, fewer special education placements, and less frequent retentions in grade at a cost that is essentially the same as that allocated to educating their control-group counterparts. The results also provide another clear, supporting example of the theoretical framework advanced by Ramey and Ramey (1998), who argue that high quality, intensive, ecologically pervasive approaches, like the schoolwide Success for All model, tend to promote meaningful sustained effects on students' academic outcomes. More generally, this study suggests that the replicable educational practices of prevention and early intervention, as modeled by Success for All, are more educationally effective, and equally expensive, relative to the traditional remedial educational practices of retention and special education.

How does the policy option of implementing Success for All compare to the policy options presented by the three other model early interventions—the Perry Preschool Program, Abecedarian Project, and the Tennessee STAR study's class-size reductions to 15 students? We explore this question in Table 4 by comparing the per-pupil costs and reading and math

achievement outcomes from the present study to the costs and reading and math outcomes of the three other interventions.[7] The average annual per-pupil cost for Success for All, across all years and schools, is multiplied by the average years of participation in the program, 3.81 years, to arrive at a total per-pupil cost. The sustained reading and math effect sizes are taken directly from our results previously shown in Table 3.

The reading and math effect sizes through age 14 of the two-year Perry Preschool program is taken from Schweinhart, Barnes, & Weikart (1993) and the effects through age 15 of the 4.5 year Abecedarian Project are from Ramey et al. (2000). The annual per-pupil expenditures associated with replicating the Abecedarian and Perry Preschool efforts rely on previously reported marginal cost estimates provided by the Developmental Center for Handicapped Persons (1987) and Barnett (1992), respectively, and are converted to constant 2000 dollars.[8] For the Tennessee STAR study, we show the independent estimates from Finn et al. (2001) and Nye et al. (1999) of the sustained effects through eighth grade of four years of exposure, from kindergarten through third grade, to reduced class sizes.[9] The current cost of replicating the STAR study's class-size reductions was determined by expressing in constant 2000 dollars the recent 1998-99 per-pupil estimate of $981 by Brewer, Krop, Gill, & Reichardt (1999).[10] The estimate provided by these authors indicates the average costs throughout the United States of reducing class sizes to15 students, but does not take into account potential costs associated with higher salaries due to greater demand for teachers and additional expenses for expanding available classroom facilities.[11] The tabulated results, thus, provide direct comparisons of the current marginal per-pupil costs of replicating each of the four model interventions along with the long-term reading and math achievement outcomes attained by the participants from the original demonstration projects at a similar point during mid-adolescence.

Although all but one of the reading and math effect estimates are relatively larger than the respective long-term Success for All effect sizes, the other interventions are also more expensive. When considering the costs and sustained effects together, Success for All provides the strongest educational benefits for the dollar for reading. For each $1,000 per-pupil expenditure, Success for All produced an effect size of 0.12, the Tennessee STAR class size reductions produced effects of 0.06 to 0.08, Perry Preschool yielded an effect of 0.03, and Abecedarian produced an effect size of 0.01. With respect to math achievement, the effect size reported by Finn et al. (2001) suggests that the Tennessee STAR and Success for All interventions produced equivalent effects per $1,000 per-pupil expenditure, and the STAR effect estimate from the Nye et al. (1999) study indicates a stronger cost-effectiveness ratio relative to Success for All. These results provide a helpful context for weighing Success for All against other policy options and other highly regarded educational interventions.[12]

Table 4
Per-Pupil Expenditures and Sustained Effects for Four Educational Interventions

	Annual PPE[a]	Years of Intervention	Total PPE[a]	Sustained Effect Size		Effect per $1000[a]	
				Reading	Math	Reading	Math
Success for All	$612	3.81	$2,331.72	0.29	0.11	0.12	0.05
Class Size Reduction to 15							
Nye et al.	$998	4.00	$3,992.00	0.32	0.37	0.08	0.09
Finn et al.	$998	4.00	$3,992.00	0.22	0.21	0.06	0.05
Perry Preschool	$8929	2.00	$17,858.00	0.51	0.49	0.03	0.03
Abecedarian Project Preschool	$10,496	4.50	$47,232.00	0.53	0.65	0.01	0.01

Note. PPE = per-pupil expenditure.
[a]All expenditure data were converted to constant 2000 dollars using gross domestic product implicit price deflators.

Another key policy issue to consider is the large-scale replicability of the programs and their effects. Success for All and Perry Preschool are the two interventions of the four that are available as nationally disseminated models. Studies from diverse localities suggest that the educational effects of the original Success for All pilot programs tend to be replicated with a good deal of consistency, but that these effects depend on the quality of the implementation (Slavin & Madden, 2001). The overall quality of implementation, though, clearly is helped by the Success for All Foundation's growing national infrastructure for supporting schools that adopt the model and by recent federal policy changes, which make more supplemental resources available for schoolwide programs like Success for All.

Similarly, the educational approach used in the Perry Preschool classrooms and home visits is widely implemented today, primarily through the use of federal Head Start funds, as the High/Scope Curriculum (Epstein, 1993). Unfortunately, though, the significant resources necessary to replicate the Perry Preschool program, as it was originally designed in Ypsilanti, typically have not been available through public programs (Kagan, 1991; Barnett, 1995). Until greater public or private commitment emerges to replicate more faithfully the intensive and relatively costly Perry Preschool model, the national replicability of the pilot program's effects remains somewhat unclear.

Widespread efforts to deliver the Abecedarian model of highly intensive health, educational, and social services to children beginning shortly after birth have not been attempted. Indeed, it is unlikely that such a program could be successfully delivered to large numbers of infants and families. These efforts would require considerable monetary investments and capacity-building efforts to establish a viable program delivery network. Once established, the program would, most likely, require considerable outreach to convince high-risk families that such invasive and intensive services were in their and their children's best interests. Therefore, although the program shows considerable promise, practical issues have limited its replication.

In recent years, the federal government has made available billions of dollars to reduce class sizes in the early grades. State-led efforts, such as California's massive initiative, also have begun recently. At least two noteworthy differences, though, set apart the Tennessee STAR model from these national and state-level initiatives. First, the Tennessee STAR class-size reductions occurred in only those schools that had the facilities to accommodate the new classrooms needed to reduce class sizes. Second, the experiment operated in a relatively small number of schools and, therefore, did not create tremendous demands for new teachers. As suggested by California's recent state-wide initiative, scaling up class-size reductions to larger numbers of schools has resulted in higher than anticipated costs, shortages of classroom space and qualified teachers, and smaller than anticipated

achievement effects (Bohrnstedt & Stecher, 1999). In addition, rather than improving equality of opportunity, Bohrnstedt and Stecher report that the California effort has exacerbated disparities between districts serving many minority and poor students and districts serving few minority and poor students.

The final point of comparison among these four interventions focuses on the research designs of the supporting studies. Because this study, unlike the other three studies of Abecedarian, Perry Preschool, and Tennessee STAR, used a quasi-experimental rather than an experimental design, causal inferences about Success for All are somewhat more tentative. The findings from this quasi-experimental study, though, are unusually robust, in that the pretest showed a statistically significant difference favoring control students but the posttest showed a statistically significant difference favoring Success for All students. This posttest difference in reading achievement was statistically significant in the ANCOVA analyses and remained statistically significant even when analyzed as a simple, unadjusted mean posttest difference for both the total sample, $t(1308)=2.74, p<.01$ (two-tailed), and low-achieving subsample, $t(330)=2.61$, $p<.01$ (two-tailed). Bracht and Glass (1968) noted the desirability of basing causal inferences on interaction patterns like this, which result in a switching of mean differences. Cook and Campbell (1979) also argued that these results are generally more conclusive than any other outcome from a no-treatment control group design because of the implausibility of alternative explanations due to scaling differences, "ceiling" effects, regression to the mean, or selection-maturation problems. Therefore, the fact that this study used a quasi-experimental design while the others employed true experimental designs may be of less significance than one might typically expect.

It is the unfortunate reality that limited funds force policymakers and practitioners to choose between educational interventions. It is tempting, therefore, to compare these model programs on the basis of their cost-effectiveness, replicability, and general strength of their supporting research and draw a summative conclusion regarding the efficacy of one over the others. However, Ramey and Ramey's (1998) principle of environmental maintenance of development suggests that such a choice of one program over another or a reliance on, for instance, only preschool intervention without elementary school and later school-based programs is misguided. Indeed, as Ramey and Ramey point out, no developmental theory is based on the assumption that positive early learning experiences are alone sufficient to ensure that children perform well throughout their lives. Our results and the results from the other model interventions consistently support these ideas.

At mid-adolescence, all students who participated in the four model interventions enjoyed outcomes that were superior to those of controls but, relative to the initial effects of the programs, these advantages had tended to wane over time, and the participants had not

generally attained normative academic outcomes. For instance, for the mid-adolescent achievement outcomes, Campbell and Ramey (1995) reported that Perry Preschool attendees scored between the 15th and 17th percentiles and control students scored below the 10th percentile. Abecedarian Project children scored between the 38th and 41st percentiles and controls scored between the 28th and 30th percentiles. The Success for All students from our analysis had eighth-grade reading and math percentile scores of 20 and 17, respectively, and controls scored at the 14th and 15th percentiles, respectively.

It is not likely that any one of these interventions could serve as the "great equalizer," or as the educational equivalent to the polio vaccine, which provides a child with protection for a lifetime all in one early dose. To compensate for poor schools, suboptimal health care, economic hardship, and other contextual conditions known to have adverse effects on at-risk students' development, educational interventions must be more akin to flu shots, which are administered throughout one's life as new risks arise within the environment. Rather than choosing one intervention over the others, the best policy may be to expand high-quality preschool programs, to continue scaling up elementary-school-based class-size reduction initiatives and Success for All programs, and to implement additional middle school and high school interventions to sustain the effects of the earlier programs.

REFERENCES

Barnett, W.S. (1985). Benefit-cost analysis of the Perry Preschool Program and its policy implications. *Educational Evaluation and Policy Analysis, 7,* 333-342.

Barnett, W.S. (1992). Benefits of compensatory preschool education. *Journal of Human Resources, 27,* 279-312.

Barnett, W.S. (1995). Long-term effects of early childhood programs on cognitive and school outcomes. *The Future of Children, 5*(3), 25-50.

Bergstralh, E.J., & Kosanke, J.L. (1995). *Computerized matching of cases to controls* (Technical Rep Ser No. 56). Rochester, MN: Section of Biostatistics, Mayo Clinic.

Bloom, B.S. (1964). *Stability and change in human characteristics.* New York: JohnWiley & Sons.

Borman, G.D., & D'Agostino, J.V. (2001). Title I and student achievement: A quantitative synthesis. In G.D. Borman, S.C. Stringfield, & R.E. Slavin (Eds.), *Title I: Compensatory education at the crossroads* (pp. 25-57). Mahwah, NJ: Erlbaum.

Bohrnstedt, G.W., & Stecher, B.M. (Eds.) (1999). *Class-size reduction in California: Early evaluation findings, 1996-1998.* (CSR Research Consortium, Year 1 Evaluation Report). Palo Alto, CA: American Institutes for Research.

Bracht, G.H., & Glass, G.V. (1968). The external validity of experiments. *American Educational Research Journal, 5,* 437-474.

Brewer, D. J., Krop, C., Gill, B.P., & Reichardt, R. (1999). Estimating the cost of national class size reductions under different policy alternatives. *Educational Evaluation and Policy Analysis, 21,* 179-192.

Campbell, F.A. & Ramey, C.T. (1995). Cognitive and school outcomes for high-risk African-American students at middle adolescence: Positive effects of early intervention. *American Educational Research Journal, 32,* 743-772.

Carter, L.F. (1984). The Sustaining Effects Study of compensatory and elementary education. *Educational Researcher, 13* (7), 4-13.

Cook, T.D., & Campbell, D.T. (1979). Quasi-experimentation: Design and analysis issues for field settings. Boston: Houghton-Mifflin.

Developmental Center for Handicapped Persons (1987). *Early Intervention Research Institute, final report.* Logan, UT: Utah State University, Author. (ERIC Document Reproduction Service No. ED293242).

Epstein, A.S. (1993). *Training for quality: Improving early childhood programs through systematic inservice training.* Ypsilanti, MI: High/Scope Press.

Finn, J.D. & Achilles, C.M. (1999). Tennessee's class size study: Findings, implications, misconceptions. *Educational Evaluation and Policy Analysis, 21,* 97-109.

Finn, J.D., Fulton, B.D., Zaharias, J., & Nye, B.A. (1989). Carry-over effects of small classes. *Peabody Journal of Education, 67,* 75-84.

Finn, J.D., Gerber, S.B., Achilles, C.M., & Boyd-Zaharias, J. (2001). The enduring effects of small classes. *Teachers College Record, 103,* 145-183.

Gomby, D.S., Larner, M.B., Stevenson, C.S., Lewit, E.M., & Behrman, R.E. (1995). Long-term outcomes of early childhood programs: Analysis and recommendations. *Future of Children, 5*(3), 6-24.

Herman, R., Aladjem, D., McMahon, P., Masem, E., Mulligan, I., O'Malley, A.S., Quinones, S., Reeve, A., & Woodruff, D. (1999). *An educator's guide to schoolwide reform.* Arlington, VA: Educational Research Service.

Jencks, C., & Phillips, M. (1998). *The black-white test score gap.* Washington, DC: Brookings.

Jones, E.M., Gottfredson, G.D., & Gottfredson, D.C. (1997). Success for some: An evaluation of a Success for All program. *Evaluation Review, 21,* 643-670.

Kagan, S.L. (1991). Excellence in early childhood education: Defining characteristics and next-decade strategies. In S.L. Kagan (Ed.), *The care and education of America's young children: Obstacles and opportunities. Ninetieth yearbook of the National Society for the Study of Education* (pp. 237-258). Chicago, IL: National Society for the Study of Education.

King, J.A. (1994). Meeting the educational needs of at-risk students: A cost analysis of three models. *Educational Evaluation and Policy Analysis, 16,* 1-19.

Lee, V.E., Brooks-Gunn, J., & Schnur, E. (1988). Are Head Start effects sustained? A longitudinal follow-up comparison of disadvantaged children attending Head Start, no preschool, and other preschool programs. *Child Development, 61,* 495-507.

Levin, H.M., Glass, G.V., & Meister, G.R. (1987). Cost-effectiveness of computer-assisted instruction. *Evaluation Review, 11,* 50-72.

Levin, H.M., & McEwan, P.J. (2001). *Cost-effectiveness analysis: Methods and applications* (2nd ed.). Thousand Oaks, CA: Sage.

McKey, R.H., Condelli, L., Ganson, H., Barret, B.J., McConkey, C., & Plantz, M.C. (1985). *The impact of Head Start on children, families, and communities.* Washington, DC: U.S. Government Printing Office.

Madden, N.A., Slavin, R.E., Karweit, N.L., Dolan, L.J., & Wasik, B.A. (1993). Success for All: Longitudinal effects of a restructuring program for inner-city elementary schools. *American Educational Research Journal, 30,* 123-148.

Nye, B., Hedges, L.V., & Konstantopoulos, S. (1999). The long-term effects of small classes: A five-year follow-up of the Tennessee class size experiment. *Educational Evaluation and Policy Analysis, 21,* 127-143.

Odden, A., & Archibald, S. (2000). *Reallocate resources: How to boost student achievement without asking for more.* Thousand Oaks, CA: Corwin.

Ramey, C.T., Campbell, F.A., Burchinal, M., Skinner, M.L., Gardner, D.M., & Ramey, S.L. (2000). Persistent effects of early childhood education on high-risk children and their mothers. *Applied Developmental Science, 4,* 2-14.

Ramey, C.T., & Ramey, S.L. (1998). Early intervention and early experience. *American Psychologist, 53,* 109-120.

Ritter, G.W., & Boruch, R.F. (1999). The political and institutional origins of a randomized controlled trial on elementary school class size: Tennessee's Project STAR. *Educational Evaluation and Policy Analysis, 21.* 111-125.

Rosenbaum, P.R. (1989). Optimal matching for observational studies. *Journal of the American Statistical Association, 84,* 1024-1032.

Schweinhart, L.J., Barnes, H.V., & Weikart, D.P. (1993). *Significant benefits: The High/Scope Perry Preschool Study through age 27.* Ypsilanti, MI: High/Scope Press.

Schweinhart, L.J., & Weikart, D.P. (1986). The effects of the Perry Preschool Program on youths through age 15 - A Summary. In Consortium for Longitudinal Studies (Ed.), *As the twig is bent...lasting effects of preschool programs* (pp. 71-101). Hillsdale, NJ: Erlbaum.

Slavin, R.E. & Madden, N.A. (with Chambers, B., Daniels, C., French, M., Haxby, B., Simons, K., & Wasik, B.) (2001). *One million children: Success for All.* Thousand Oaks, CA: Corwin.

Slavin, R.E., Madden, N.A., Dolan, L.J., Wasik, B.A., Ross, S.M., & Smith, L.M. (1994). 'Whenever and wherever we choose:' The replication of 'Success for All. *Phi Delta Kappan, 75,* 639-647.

Slavin, R.E., Madden, N.A., Karweit, N.L., Dolan, L.J., & Wasik, B.A. (1990). *Success for All: effects of variations in duration and resources of a schoolwide elementary restructuring program* (Report No. 2). Baltimore, MD: Johns Hopkins University, Center for Research on Effective Schooling for Disadvantaged Students.

Slavin, R.E., Madden, N.A., Karweit, N.L., Dolan, L.J., & Wasik, B.A. (1992). *Success for All: A relentless approach to prevention and early intervention in elementary schools.* Arlington, VA: Educational Research Service.

Slavin, R.E., Madden, N.A., Karweit, N.L., Livermon, B.J., & Dolan, L.J. (1990). Success for All: First-year outcomes of a comprehensive plan for reforming urban education. *American Educational Research Journal, 27,* 255-278.

Stasz, C., & Stecher, B.M. (2000). Teaching mathematics and language arts in reduced size and non-reduced size classrooms. *Educational Evaluation and Policy Analysis, 22,* 313-329.

U.S. Bureau of the Census (1997). *Public education finances: 1996-1997*. Washington, DC: U.S. Government Printing Office.

U.S. Department of Labor, Bureau of Labor Statistics (1999). *Occupational employment and wage estimates, 1998*. [On-line]. Available: http://stats.bls.gov/oeshome.htm.

White, K.R. (1985). Efficacy of early intervention. *Journal of Special Education, 19*, 401-416.

NOTES

[1] Finn et al. (2001) and Nye et al. (1999) both analyzed the followup data to the Tennessee STAR study and arrived at somewhat different sustained effect estimates for students who were taught in small classes during all four years of the study. We present the outcomes reported in both studies and suggest that they should be interpreted as an estimated range of potential effects.

[2] Compare, for example, the original Success for All pilot program as described by Slavin, Madden, Karweit, Dolan, & Wasik, (1990) and a more recent description by Slavin and Madden (2001) of the current model that is available.

[3] The added resources for implementing Success for All came from various sources. The funding for the Abbottston Success for All program was provided by Chapter 2 funds, City Springs paid for its implementation through a grant from a private foundation, and the remaining Success for All schools used funds from a U.S. Department of Education dropout prevention grant (Slavin, Madden, Karweit, Dolan, & Wasik, 1990).

For three schools, Dallas Nicholas, Bernard Harris, and Harriet Tubman, the expenses related to tutors were, essentially, "costless" relative to the control schools as these schools reassigned their Title I reading instructors as Success for All tutors. The control schools also had Title I reading teachers, but in the control schools, the Title I instructors continued teaching traditional reading pullout classes rather than offering tutoring help, as specified by the Success for All model. The difference between the Success for All costs when one considers this a costless redeployment, as we have, and the Success for All costs when one considers this redeployment as an added cost is expressed by the difference between our cost figures in Table 2 labeled, respectively, marginal costs and total costs.

[4] Although the Success for All and control student attrition rates were statistically equivalent, there is a possibility that the Success for All students who dropped out of our analyses were systematically different from the control students who dropped from our analyses. Such differences could bias our estimates of the program effects and compromise the internal validity of the study. In addition, it is possible that those Success for All and control students who dropped from our analyses were systematically different from those who remained in the analytical samples. Differences of this sort could limit to whom we might generalize our results and, thus, might compromise the study's external validity.

First, we contrasted the background characteristics of those Success for All and control students who dropped out of our analyses. The two groups were statistically equivalent on all background characteristics, with the exception of the statistically significant difference on the reading pretest score for those who dropped out of the transcript analysis sample, $t(868)=-5.97$, $p<.001$ (two-tailed), $g=-0.40$, and the achievement test analysis sample, $t(1291)=-6.94$, $p<.001$ (two-tailed), $g=-0.38$. The magnitudes of these differences, though, were essentially the same as the magnitudes of the pretest differences for the samples that we retained for our transcript analyses ($g=-0.35$) and achievement analyses ($g=-0.35$). Therefore, these results do not provide evidence of differential attrition.

Second, with respect to the study's external validity, we contrasted the background variables of Success for All and control students who dropped out of our study to the backgrounds of Success for All and control students who we were able to retain in our analyses. The students retained in our transcript analysis samples and the students who dropped from our transcript analysis were statistically equivalent on all background variables. However, for the achievement analyses, those who were retained in the analysis had higher pretest scores than those who were not retained for both the Success for All sample, $t(1136)=-4.68$, $p<.001$ (two-tailed), and for the control sample, $t(1463)=-3.70$, $p<.001$ (two-tailed). These results suggest that our analyses of the transcript outcomes have greater external validity than our analyses of the achievement outcomes. However, given the consistency of the outcomes across the transcript and achievement analyses, it does not appear that this discrepancy had serious consequences for the general direction and magnitudes of the effect estimates.

[5] One-on-one computerized matching of Success for All and control students was attempted using an algorithm developed by Bergstralh and Kosanke (1995), which is based on the optimal matching procedures described by Rosenbaum (1989). Statistically significant Success for All-control pretest differences remained even after the optimal one-to-one Success for All-control pretest matches were identified by the computerized matching procedure.

[6] Encouragingly, the marginal cost estimates we derived based on the ingredients method were somewhat consistent with the original marginal costs estimated by Slavin, Madden, Karweit, Dolan, & Wasik (1992). Slavin et al. reported that the marginal costs, in 1988 dollars, of two of the original Success for All implementations, at Abbottston and City Springs Elementary, were $400,000 and the marginal costs of the remaining three programs were $40,000. The developer indicated that these estimates included all program ingredients, personnel, training, materials, and professional development, at Abbottston and City Springs (Robert Slavin, personal communication, 11/1/00). At the other three schools, though, existing personnel were shifted to fill most of the new positions required by the Success for All program. Marginal costs at these schools were limited to a half-time program facilitator, training, materials, and professional development.

To compare our marginal cost estimates to the original amounts indicated by Slavin et al. (1992) (i.e., $400,000 for Abbottston and City Springs and $40,000 for the remaining three Success for All schools), consider the school-by-school averages of our Year 1, 2, and 3 marginal cost estimates converted to 1988 dollars: Abbottston, $323,144.12; City Springs, $436,373.87; Dallas Nicholas, $71,117.18; Dr. Bernard Harris, $71,117.18; Harriett Tubman, $71,117.18. For every school but Abbottston, our estimates were higher than the developer's original estimates. This comparison suggests that, if anything, our method tended to provide a relatively liberal estimate of the costs of implementing Success for All.

Because varying cost assumptions may lead to different results, we conducted further sensitivity analyses that employed 1987-88 Baltimore City Public School System data on: (a) average per-pupil expenditures; (b) average per-pupil special education expenditures; and (c) personnel salaries and benefits rates (for calculating the personnel costs associated with Success for All implementation). Because no historic estimates of the Success for All

materials, training, and professional development costs were available, we used the same cost data reported on the Success for All internet site. In general, analyses of these data, adjusted to constant 2000 dollars, indicated that the costs of schooling through the successful completion of eighth grade for both control and Success for All students were somewhat less than the costs reported in Table 3. Because the Success for All-control differences, though, were essentially the same as those reported in Table 3, we elected to use the reported national cost estimates, which provide policymakers and practitioners more pertinent information regarding the national expenditures associated with Success for All and control students' schooling. The results of the sensitivity analyses are available from the first author on request.

[7] We have compared the interventions across these two outcomes, reading and math achievement, for two main reasons. First, from a practical standpoint, these are the primary outcomes that are consistently documented across the studies of all four interventions. Second, and more importantly, we consider reading and math achievement outcomes as two key educational outcomes at eighth grade. We also believe that these outcomes are strong predictors of other future educational and adult outcomes. Indeed, as Jencks and Phillips (1998) argue, reducing the test score gap is probably both necessary and sufficient for substantially reducing inequality in educational attainment and earnings. It is important to note, though, that these comparisons do not go beyond eighth grade and, therefore, exclude some data, most notably from the Perry Preschool intervention, that show lasting effects into adulthood on outcomes including social adjustment and economic success (Barnett, 1992; Schweinhart et al., 1993). These long-term benefits to students and society have considerably outweighed the substantial costs associated with the original preschool program (Barnett, 1985; Schweinhart & Weikart, 1986).

[8] All per-pupil costs cited in Table 4 represent a similar estimate of each program's marginal costs; that is, the amount spent on program students beyond that which was spent on control students. However, some may argue that the elementary school programs, Success for All and Tennessee class-size reduction, have a sort of unfair advantage over the preschool programs, in that the control students at the elementary level receive at least a base level of expenditure provided by the public school system. If we were to assume that such a base existed for the preschool control students, how would that impact our results? Using as a base the average fiscal year 2000 per-pupil expense of Head Start, which was reported by the Head Start Bureau as $5,951 (see *http://www2.acf.dhhs.gov/programs/hsb/about/fact2001.htm*), we found that the overall additional expense of the two years of Perry Preschool would be $5,956 and the added expense of the four and a half years of the Abecedarian program would be $20,452.50. Using these figures, the effect per $1,000 investment in Perry Preschool is 0.09 and the effect per $1,000 investment in Abecedarian is 0.03. This adjustment, therefore, does not fundamentally alter our conclusion regarding Success for All's overall cost-effectiveness advantage.

[9] These effects from the Tennessee STAR study are based on the select subsample of students who experienced small classes during all four years of the study. Our Success for All effect estimate is based on a more conservative treatment definition, which is akin to an "intention-to-treat" definition: that is, we analyzed the effects of initial first-grade enrollment in a Success for All school regardless of students' participation in later years. Students' years of attendance in Success for All schools and small classes, though, were similar.

[10] The estimate of $998, in constant 2000 dollars, by Brewer et al. (1999) is very similar to a basic per-pupil estimate of the costs particular to the Tennessee STAR study. Specifically, Ritter and Boruch (1999) reported that in 1985 the Tennessee legislature appropriated $3 million for each year of the 4-year study. Ritter and Boruch stated that the majority of these funds were used for hiring the additional teachers and classroom aides required by the project. Using the $3 million figure and Ritter and Boruch's estimates of 1,900 students originally assigned to small classes and 2,200 students originally assigned to classrooms with aides, we may establish a yearly per-pupil cost estimate of $609.76 in 1985 dollars, or $886.24 in constant 2000 dollars.

[11] This represents one of two important issues that may be of considerable consequence in our cost estimate for class-size reductions. In both regards, Levin, Glass, and Meister (1987) provide helpful information. First, the authors indicated that the additional costs of a classroom, which include the physical space, furnishings, energy needs, insurance, and maintenance in addition to a teacher, totaled $28,138 in 1980 dollars. Converting this amount to constant 2000 dollars using gross domestic product implicit price deflators, the cost is $54,172. If we assume an average class-size reduction in the Tennessee STAR study of 40%, from 25 to 15 students, we obtain a per-pupil expenditure estimate of $1,444.59 for a small class of 15 students (0.40 * $54,172 / 15=$1,444.59). This estimate, which includes the additional expenses required for classroom space but does not include potential increases in teacher salaries, is 31% higher than the figure reported by Brewer et al. (1999) that we used in our cost-effectiveness estimates in Table 4.

Potentially impacting the cost estimate associated with class-size reductions in the opposite direction, Levin and his colleagues argued that "reduction in class size is an overall educational intervention that should affect all of the educational activities during the school day" (p. 64). As a result, the authors argued that only a portion of the marginal cost of class-size reduction should be viewed as an educational intervention to improve a single subject area, such as reading or math. Levin et al. stated that about one third of the school day is devoted to reading, and the other two thirds is spent on other activities. Accordingly, they divided the total marginal cost of class-size reduction by three to obtain an estimated cost for reading instruction and for math instruction.

Similar to this line of reasoning, and as suggested by our theoretical framework, our perspective is that reductions in class size, Success for All, the Perry Preschool, and the Abecedarian Project all represent *comprehensive* educational programs. Accordingly, we have presented a range of educational outcomes for Success for All and have compared the relative effects of the four programs on the basis of two key educational outcomes: reading and math achievement. This latter adjustment suggested by the work of Levin at al., therefore, is not applicable to our theoretical and analytical framework, as the same adjustment could be applied to each of the four interventions we have considered.

[12] Odden and Archibald (2000) recently provided useful analyses and demonstrations of how schools may implement full-scale Success for All programs without spending any additional resources. Through resource reallocation, it may make it possible for schools to implement Success for All, to achieve similar educational outcomes to those documented here, and to pay little to no additional per-pupil expenses.

CRESPAR

Johns Hopkins University
Center for Social Organization of Schools
3003 North Charles Street – Suite 200
Baltimore MD 21218
410-516-8800 / 410-516-8890 fax

Howard University
2900 Van Ness Street, NW
Washington DC 20008
202-806-8484 / 202-806-8498 fax